

# UNIT BALLS OF CONSTANT VOLUME: WHICH ONE HAS OPTIMAL REPRESENTATION?

JEAN B. LASSERRE

**ABSTRACT.** In the family of unit balls with constant volume we look at the ones whose algebraic representation has some extremal property. We consider the family of nonnegative homogeneous polynomials of even degree  $d$  whose sublevel set  $\mathbf{G} = \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$  (a unit ball) has same fixed volume and want to find in this family the one that minimizes either the  $\ell_1$ -norm or the  $\ell_2$ -norm of its vector of coefficients. Equivalently, among all degree- $d$  polynomials of constant  $\ell_1$ - or  $\ell_2$ -norm, which one minimizes the volume of its level set  $\mathbf{G}$ . We first show that in both cases this is a convex optimization problem with a unique optimal solution  $g_1^*$  and  $g_2^*$  respectively. We also show that  $g_1^*$  is the  $L_p$ -norm polynomial  $\mathbf{x} \mapsto \sum_{i=1}^n x_i^p$ , thus recovering a parsimony property of the  $L_p$ -norm via  $\ell_1$ -norm minimization. (Indeed  $n = \|g_1^*\|_0$  is the minimum number of non-zero coefficient for  $\mathbf{G}$  to have finite volume.) This once again illustrates the power and versatility of the  $\ell_1$ -norm relaxation strategy in optimization when one searches for an optimal solution with parsimony properties. Next we show that  $g_2^*$  is not sparse at all (and so differs from  $g_1^*$ ) but is still a sum of  $p$ -powers of linear forms. We also characterize the unique optimal solution of the same problem where one searches for an SOS homogeneous polynomial that minimizes the trace of its associated (psd) Gram matrix, hence aiming at finding a solution which is a sum of a few squares only. Finally, we also extend these results to generalized homogeneous polynomials, which includes  $L_p$ -norms when  $0 < p$  is rational.

## 1. INTRODUCTION

It is well-known that the shape of the Euclidean unit ball  $\mathbf{B}_2 = \{\mathbf{x} : \sum_{i=1}^n x_i^2 \leq 1\}$  has spectacular geometric properties with respect to other shapes. For instance, the sphere has the smallest surface area among all surfaces enclosing a given volume and it encloses the largest volume among all closed surfaces with a given surface area; Hilbert and Cohn-Vossen [8] even describe eleven geometric properties of the sphere!

But  $\mathbf{B}_2$  has also another spectacular (non-geometric) property related to its algebraic representation which is obvious even to people with a little background in Mathematics: Namely, its defining polynomial  $\mathbf{x} \mapsto g_2(\mathbf{x}) := \sum_{i=1}^n x_i^2$  cannot be simpler!! Indeed, among all nonnegative quadratic homogeneous polynomials  $\mathbf{x} \mapsto g(\mathbf{x}) = \sum_{i \leq j} g_{ij} x_i x_j$  that define a bounded ball  $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ ,  $g_2$  is the one that minimizes the “cardinality norm”  $\|g\|_0 := \#\{(i, j) : g_{ij} \neq 0\}$  (which actually is not a norm). Only  $n$  coefficients of  $g_2$  do not vanish and there cannot be less than  $n$  non zero coefficients to define a bounded ball  $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ . The same is true for the  $d$ -unit ball  $\mathbf{B}_d = \{\mathbf{x} : \sum_{i=1}^n x_i^d \leq 1\}$  and its defining polynomial  $\mathbf{x} \mapsto g_d(\mathbf{x}) = \sum_i x_i^d$  for any even integer  $d$ , when compared to any other nonnegative homogeneous polynomial  $g$  of degree  $d$  whose sublevel set  $\{\mathbf{x} : g(\mathbf{x}) \leq 1\}$  has finite Lebesgue volume. Indeed, again  $\|g_d\|_0 = n$ , i.e., out of potentially  $\binom{n+d-1}{d}$

coefficients only  $n$  do not vanish! In other words,

$$(1.1) \quad g_d = \arg \min_g \{ \|g\|_0 : \text{vol}(\{\mathbf{x} : g(\mathbf{x}) \leq 1\}) \leq 1 \}$$

where the minimum is taken over all homogeneous polynomials of degree  $d$ .

So an natural question which arises is as follows: *In view of the many “geometric properties” of the unit ball  $\mathbf{B}_d$ , is the “algebraic sparsity” of its representation  $\{\mathbf{x} : \sum_i x_i^d \leq 1\}$  a coincidence or does it also corresponds to a certain extremal property on all possible representations?*

So we are interested in the following optimization problem in computational geometry and with an algebraic flavor.

*Given an even integer  $d$ , determine the nonnegative homogeneous polynomial  $g^*$  of degree  $d$  whose  $\ell_1$ -norm  $\|g^*\|_1$  (or  $\ell_2$ -norm  $\|g^*\|_2$ ) of its vector of coefficients is minimum among all degree- $d$  nonnegative homogeneous polynomials with same (fixed) volume of their sublevel set  $\mathbf{G} = \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ . That is, solve:*

$$(1.2) \quad \inf_g \{ \|g\|_{p=1,2} : \text{vol}(\mathbf{G}) = 1; g \text{ homogeneous of degree } d \}.$$

In particular, *Can the parsimony property of the  $L_d$ -unit balls be recovered from (1.2) with the  $\ell_1$ -norm  $\|g\|_1$  (instead of minimizing the nasty function  $\|\cdot\|_0$  in (1.1))?*

By homogeneity, this problem also has the equivalent formulation: *Among all homogeneous polynomials  $g$  of degree  $d$  and with constant norm  $\|g\|_1 = 1$  (or  $\|g\|_2 = 1$ ) find the one with level set  $\mathbf{G}$  of minimum volume.*

One goal of this paper is to prove that (1.2) is a convex optimization problem with a *unique optimal solution*, which is the same as  $g_d$  in (1.1) when one minimizes the  $\ell_1$ -norm  $\|g\|_1$ . In addition  $g_d$  cannot be an optimal solution of (1.2) when one minimizes the  $\ell_2$ -norm  $\|g\|_2$  (except when  $d = 2$ ). This illustrates in this context of computational geometry that again, the sparsity-induced  $\ell_1$ -norm does a perfect job in the relaxation (1.2) (with  $\|\cdot\|_1$ ) of problem (1.1) with  $\|\cdot\|_0$ . This convex “relaxation trick” in (non convex)  $\ell_0$ -optimization has been used successfully in several important applications; see e.g. Candès et al. [4], Donoho [5], Donoho and Elad [6] in compressed sensing applications and Recht et al. [14] for matrix applications (where the small-rank induced *nuclear norm* is the matrix analogue of the  $\ell_1$ -norm). For more details on optimization with sparsity constraints and/or sparsity-induced penalties, the interested reader is referred to Beck and Eldar [3] and Bach et al. [2].

To address our problem we consider the following framework: Let  $\mathbf{Hom}_d \subset \mathbb{R}[\mathbf{x}]_d$  be the vector space of homogeneous polynomials of even degree  $d$ , and given  $g \in \mathbf{Hom}_d$ , let  $\mathbf{g} = (g_\alpha)$  be its vector of coefficients, i.e.,

$$\mathbf{x} \mapsto g(\mathbf{x}) = \sum_{\alpha} g_{\alpha} \mathbf{x}^{\alpha} \left( = \sum_{\alpha} g_{\alpha} \mathbf{x}_1^{\alpha_1} \cdots \mathbf{x}_n^{\alpha_n} \right), \quad \sum_i \alpha_i = d,$$

with standard  $\ell_1$ -norm  $\|g\|_1 = |\mathbf{g}| = \sum_{\alpha} |g_{\alpha}|$ . With any  $g \in \mathbf{Hom}_d$  is associated its sublevel set  $\mathbf{G} \subset \mathbb{R}^n$  defined by:

$$(1.3) \quad \mathbf{G} := \{ \mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \leq 1 \}, \quad g \in \mathbf{Hom}_d.$$

In particular, with  $\mathbf{x} \mapsto g^*(\mathbf{x}) := \sum_{i=1}^n x_i^d$ , the sublevel set  $\mathbf{G}^*$  is nothing less than the standard  $d$ -unit ball

$$\mathbf{B}_d = \{ \mathbf{x} : \sum_{i=1}^n x_i^d \leq 1 \} =: \{ \mathbf{x} : \|\mathbf{x}\|_d^d \leq 1 \},$$

whose Lebesgue volume  $\text{vol}(\mathbf{B}_d)$  is denoted  $\rho_d$ . (When  $g \in \mathbf{Hom}_d$  is convex then  $\mathbf{x} \mapsto g(\mathbf{x})$  defines a norm  $\|\mathbf{x}\|_g := g(\mathbf{x})^{1/d}$  with  $\mathbf{G}$  as associated unit ball.)

**Contribution.** (a) In a first contribution we prove that the optimization problem:

$$(1.4) \quad \mathbf{P}_1 : \inf_g \{ \|g\|_1 : \text{vol}(\mathbf{G}) \leq \rho_d; \quad g \in \mathbf{Hom}_d \},$$

has a unique optimal solution  $g_1^*$  which is the  $L_d$ -norm polynomial  $\mathbf{x} \mapsto \|\mathbf{x}\|_d^d$ . Observe that  $g_1^*$  has the minimal number  $n$  of coefficients over potentially  $s(d) := \binom{n-1+d}{d}$  coefficients. (Indeed for a polynomial  $g \in \mathbf{Hom}_d$  with  $m < n$  non zero coefficients, its sublevel set  $\mathbf{G}$  cannot have finite Lebesgue volume.) Therefore the  $L_d$ -norm polynomial  $g_1^*$  associated with the unit ball  $\mathbf{B}_d$  is the “sparsest” solution among all  $g \in \mathbf{Hom}_d$  such that  $\text{vol}(\mathbf{G}) \leq \text{vol}(\mathbf{B}_d)$ . In particular,  $g_1^*$  not only solves problem  $\mathbf{P}_1$  but also solves the non convex optimization problem

$$\mathbf{P}_0 : \inf_g \{ \|g\|_0 : \text{vol}(\mathbf{G}) \leq \rho_d; \quad g \in \mathbf{Hom}_d \},$$

of which  $\mathbf{P}_1$  is a “convex relaxation”. But this is also equivalent to state that among all homogeneous polynomials of degree  $d$  with constant  $\ell_1$ -norm, the  $L_d$ -unit ball is the one with minimum volume  $\text{vol}(\mathbf{B}_d) \leq \text{vol}(\mathbf{G})$ .

(b) In a second contribution we consider the  $\ell_2$ -norm version of (1.4):

$$(1.5) \quad \mathbf{P}_2 : \inf_g \{ \|g\|_2 : \text{vol}(\mathbf{G}) \leq \rho_d; \quad g \in \mathbf{Hom}_d \},$$

with weighted Euclidean norm  $g \mapsto \|g\|_2$  defined by:

$$\|g\|_2^2 := \sum_{|\alpha|=d} c_\alpha g_\alpha^2, \quad g \in \mathbf{Hom}_d, \quad \text{where } c_\alpha := \frac{(d)!}{\alpha_1! \cdots \alpha_n!}.$$

We then show that  $\mathbf{P}_2$  also has a unique optimal solution  $g_2^*$ , but in contrast to the optimal solution  $g_1^*$  of problem  $\mathbf{P}_1$ ,  $g_2^*$  is not sparse at all! This is because one can show that all  $\binom{n-1+d}{d}$  coefficients of the form  $g_{2\beta}^*$  with  $|\beta| = d/2$  are non-zero. In addition,  $g_2^*$  is a particular sum of squares (SOS) polynomial as it is a sum of  $d$ -powers of linear forms. (Notice that  $g_1^*$  is also a (very particular and simple) sum of  $d$ -powers of linear forms.) In particular, when  $d = 4$  the optimal solution of  $\mathbf{P}_2$  is the Euclidean ball  $\{ \mathbf{x} : \sum_i x_i^2 \leq 1 \}$  which has the equivalent quartic representation  $\{ \mathbf{x} : (\sum_{i=1}^n x_i^2)^2 \leq 1 \}$  and the SOS quartic polynomial  $\mathbf{x} \mapsto (\sum_{i=1}^n x_i^2)^2$  solves  $\mathbf{P}_2$ .

(c) We also consider the SOS (sum of squares) version of  $\mathbf{P}_1$ , that is one now searches for a degree- $d$  SOS homogeneous polynomial  $g_{\mathbf{Q}}(\mathbf{x}) = \mathbf{v}_{d/2}(\mathbf{x}) \mathbf{Q} \mathbf{v}_{d/2}(\mathbf{x})$ ,  $\mathbf{Q} \succeq 0$ , (where  $\mathbf{v}_{d/2}(\mathbf{x}) = (\mathbf{x}^\alpha)$ ,  $|\alpha| = d/2$ ). That is, one characterizes the unique optimal solution of the optimization problem:

$$(1.6) \quad \mathbf{P}_3 : \inf_{\mathbf{Q} \succeq 0} \{ \text{trace}(\mathbf{Q}) : \text{vol}(\mathbf{G}_{\mathbf{Q}}) \leq \rho_d; \quad \mathbf{Q} \succeq 0 \}.$$

In this matrix context,  $\text{trace}(\mathbf{Q})$  is the *nuclear norm* of  $\mathbf{Q}$  and so solving  $\mathbf{P}_3$  aims at finding an optimal solution  $\mathbf{Q}^*$  with small rank, which translates into an homogeneous polynomial  $g_{\mathbf{Q}^*}$  which is a sum of a few squares. We also prove that  $g_1^*$

associated with the  $L_d$ -unit ball cannot be an optimal solution of  $\mathbf{P}_3$  (and indeed  $g_{\mathbf{Q}^*}$  being a sum of a few squares does not necessarily implies that it has a small number of coefficients).

(d) Finally we also show that results in (a) and (b) extend to the case of other values of  $d$  (including  $p = 1$  and rationals) in which case one now deals with positively homogeneous “generalized polynomials” (instead of homogeneous polynomials) and one has to define an appropriate finite-dimensional analogue analogue of  $\mathbf{Hom}_d$ . This includes the interesting case of the  $L_1$ -unit ball  $\{\mathbf{x} : \sum_i |x_i| \leq 1\}$  and when  $p < 1$ , balls which are not associated with norms.

## 2. NOTATION, DEFINITIONS AND PRELIMINARY RESULTS

Let  $\mathbb{R}[\mathbf{x}]$  denote the ring of real polynomials in the variables  $\mathbf{x} = (x_1, \dots, x_n)$ , and let  $\mathbb{R}[\mathbf{x}]_d$  be the vector space of real polynomials of degree at most  $d$ . Similarly, let  $\Sigma[\mathbf{x}] \subset \mathbb{R}[\mathbf{x}]$  denote the convex cone of real polynomials that are sums of squares (SOS) of polynomials, and  $\Sigma[\mathbf{x}]_d \subset \Sigma[\mathbf{x}]$  its subcone of SOS polynomials of degree at most  $d$ . Denote by  $\mathcal{S}^m$  the space of  $m \times m$  real symmetric matrices. For a given matrix  $\mathbf{A} \in \mathcal{S}^m$ , the notation  $\mathbf{A} \succeq 0$  (resp.  $\mathbf{A} \succ 0$ ) means that  $\mathbf{A}$  is positive semidefinite (psd) (resp. positive definite (pd)), i.e., all its eigenvalues are real and nonnegative (resp. positive).

A polynomial  $p \in \mathbb{R}[\mathbf{x}]_d$  is homogenous if  $p(\lambda \mathbf{x}) = \lambda^d p(\mathbf{x})$  for all  $\mathbf{x} \in \mathbb{R}^n$ ,  $\lambda \in \mathbb{R}$ . A function  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is positively homogeneous of degree  $d \in \mathbb{R}$  if  $f(\lambda \mathbf{x}) = \lambda^d f(\mathbf{x})$  for all  $0 \neq \mathbf{x} \in \mathbb{R}^n$ ,  $\lambda > 0$ . For instance  $x \mapsto |x|$  is not homogeneous but is positively homogeneous of degree 1.

Let  $\mathbf{Hom}_d \subset \mathbb{R}[\mathbf{x}]_d$  be the vector space of homogeneous polynomials of even degree  $d$ , and let  $\mathbb{N}_d^n := \{(\alpha_1, \dots, \alpha_n) : \sum_i \alpha_i = d\}$ . For an homogeneous polynomial  $g \in \mathbb{R}[\mathbf{x}]_d$ , and with  $s(d) := \binom{n-1+d}{d}$ , let  $\mathbf{g} = (g_\alpha) \in \mathbb{R}^{s(d)}$  be its vector of coefficients, i.e.,

$$\mathbf{x} \mapsto g(\mathbf{x}) := \sum_{\alpha \in \mathbb{N}_d^n} g_\alpha \mathbf{x}^\alpha \left( = \sum_{\alpha \in \mathbb{N}_d^n} g_\alpha x_1^{\alpha_1} \cdots x_n^{\alpha_n} \right).$$

Denote by  $\mathbf{G} \subset \mathbb{R}^n$  its associated sublevel set  $\mathbf{G} := \{\mathbf{x} : g(\mathbf{x}) \leq 1\}$ .

Let  $\mathbf{P}[\mathbf{x}]_d \subset \mathbf{Hom}_d$  be the set of homogeneous polynomials of degree  $d$  whose associated level set  $\mathbf{G}$  has finite Lebesgue volume. It is a convex cone; see [10, Proposition 2.1]. Let  $f : \mathbf{Hom}_d \rightarrow \mathbb{R}_+$  be the function defined by:

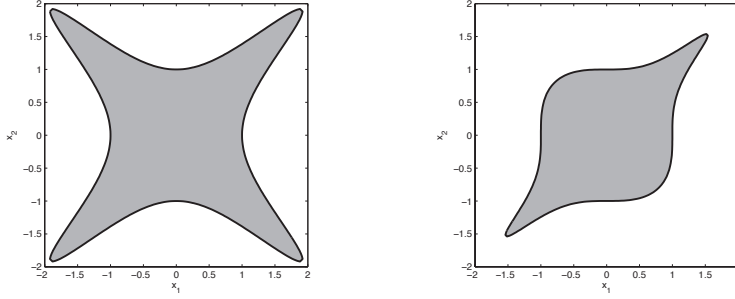
$$g \mapsto f(g) := \begin{cases} \text{vol}(\mathbf{G}) & \text{if } g \in \mathbf{P}[\mathbf{x}]_d \\ +\infty & \text{otherwise.} \end{cases}$$

It is important to realize that the sublevel set  $\mathbf{G}$  need not be convex! For instance Figure 2 displays two examples of non convex sets  $\mathbf{G}$ .

More generally, for every  $\alpha \in \mathbb{N}^n$  define  $f_\alpha : \mathbf{P}[\mathbf{x}]_d \rightarrow \mathbb{R}$  by

$$g \mapsto f_\alpha(g) := \begin{cases} \int_{\mathbf{G}} \mathbf{x}^\alpha d\mathbf{x} & \text{if } g \in \mathbf{P}[\mathbf{x}]_d \\ +\infty & \text{otherwise.} \end{cases}$$

In particular,  $f(g) = f_0(g)$ .

FIGURE 1.  $\mathbf{G}_1$  with  $x^4 + y^4 - 1.925 x^2 y^2$  and  $x^6 + y^6 - 1.925 x^3 y^3$ 

**A preliminary result.** We will need the following result of independent interest already proved in [10] but for which we provide a brief sketch. We use the same technique based on Laplace transform as in Lasserre [11] and Lasserre and Zeron [13] for providing closed form expressions for certain class of integrals.

**Theorem 2.1.** *Let  $h : \mathbb{R}^n \rightarrow \mathbb{R}$  be a nonnegative positively homogeneous function of degree  $0 \neq d \in \mathbb{R}$  such that  $\text{vol}(\{\mathbf{x} : h(\mathbf{x}) \leq 1\}) < \infty$ . Then for every  $\alpha \in \mathbb{N}^n$ :*

$$(2.1) \quad \int_{\{\mathbf{x}: h(\mathbf{x}) \leq 1\}} \mathbf{x}^\alpha d\mathbf{x} = \frac{1}{\Gamma(1 + (n + |\alpha|)/d)} \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-h(\mathbf{x})) d\mathbf{x}.$$

In particular when  $d$  is an even integer: For every  $g \in \mathbf{P}[\mathbf{x}]_d$ ,

$$(2.2) \quad f(g) = \text{vol}(\mathbf{G}) = \frac{1}{\Gamma(1 + n/d)} \int_{\mathbb{R}^n} \exp(-g(\mathbf{x})) d\mathbf{x},$$

and the function  $f$  is nonnegative, strictly convex and homogeneous of degree  $-n/d$ . Moreover, if  $g \in \text{int}(\mathbf{P}[\mathbf{x}]_d)$ :

$$(2.3) \quad \frac{\partial f(g)}{\partial g_\alpha} = -\frac{n+d}{d} \int_{\mathbf{G}} \mathbf{x}^\alpha d\mathbf{x}, \quad \alpha \in \mathbb{N}_d^n$$

$$(2.4) \quad \int_{\mathbf{G}} g(\mathbf{x}) d\mathbf{x} = \frac{n}{n+d} \int_{\mathbf{G}} d\mathbf{x}.$$

*Proof.* For  $\alpha \in \mathbb{N}^n$ , let  $v_\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}$  be the function  $y \mapsto v_\alpha(y) := \int_{\{\mathbf{x}: h(\mathbf{x}) \leq y\}} \mathbf{x}^\alpha d\mathbf{x}$ . Observe that  $v_\alpha(y) = 0$  whenever  $y < 0$ . So let  $\mathcal{L}[v_\alpha] : \mathbb{C} \rightarrow \mathbb{C}$  be the Laplace transform  $\mathcal{L}[v_\alpha]$  of the function  $v_\alpha$ , i.e.,

$$\lambda \mapsto \mathcal{L}[v_\alpha](\lambda) := \int_0^\infty \exp(-\lambda y) v_\alpha(y) dy, \quad \lambda \in \mathbb{C}; \Re(\lambda) > 0.$$

Let  $H_\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}$  be the function  $\lambda \mapsto H_\alpha(\lambda) := \mathcal{L}[v_\alpha](\lambda)$ ,  $\lambda \in \mathbb{R}_+$ . Then:

$$\begin{aligned}
H_\alpha(\lambda) &= \int_0^\infty \exp(-\lambda y) \left( \int_{\{\mathbf{x}: h(\mathbf{x}) \leq y\}} \mathbf{x}^\alpha d\mathbf{x} \right) dy \\
&= \int_{\mathbb{R}^n} \mathbf{x}^\alpha \left( \int_{h(\mathbf{x})}^\infty \exp(-\lambda y) dy \right) d\mathbf{x} \\
&= \frac{1}{\lambda} \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-\lambda h(\mathbf{x})) d\mathbf{x} \\
&= \frac{1}{\lambda^{1+(n+|\alpha|)/d}} \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-h(\mathbf{x})) d\mathbf{x} \quad [\text{by homogeneity}] \\
&= \frac{\Gamma(1+(n+|\alpha|)/d)}{\lambda^{1+(n+|\alpha|)/d}} \underbrace{\left( \frac{1}{\Gamma(1+(n+|\alpha|)/d)} \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-h(\mathbf{x})) d\mathbf{x} \right)}_c \\
&= c \frac{\Gamma(1+(n+|\alpha|)/d)}{\lambda^{1+(n+|\alpha|)/d}}.
\end{aligned}$$

The function  $H_\alpha$  is analytic on  $\Theta = \{\lambda \in \mathbb{C} : \Re(\lambda) > 0\}$  and coincides with  $\mathcal{L}[v_\alpha]$  on  $[0, \infty)$ . By the identity Theorem on analytic functions (see e.g. [7, Theorem III.3.2, p. 125]),  $H_\alpha = \mathcal{L}[v_\alpha]$  on  $\Theta$ . But  $H_\alpha = \mathcal{L}[c y^{(n+|\alpha|)/d}]$  which concludes the proof and yields (2.1) when  $y = 1$ .

Next when  $g \in \text{int}(\mathbf{P}[\mathbf{x}]_d)$ , one obtains (2.3) by differentiating under the integral sign which is permitted in this context; see [10] for a rigorous proof. This yields

$$\begin{aligned}
\frac{\partial f(g)}{\partial g_\alpha} &= \frac{-1}{\Gamma(1+n/d)} \int_{\mathbb{R}^n} \mathbf{x}^\alpha \exp(-g(\mathbf{x})) d\mathbf{x}, \quad \alpha \in \mathbb{N}_d^n \\
&= \frac{-\Gamma(1+(n+d)/d)}{\Gamma(1+n/d)} \int_{\mathbf{G}} \mathbf{x}^\alpha d\mathbf{x}, \quad \alpha \in \mathbb{N}_d^n \\
&= -\frac{n+d}{d} \int_{\mathbf{G}} \mathbf{x}^\alpha d\mathbf{x}, \quad \alpha \in \mathbb{N}_d^n,
\end{aligned}$$

where we have use the identity  $\Gamma(z+1) = z\Gamma(z)$ . Finally, to get (2.4) observe that  $f$  is a positively homogeneous function of degree  $-n/d$  and so Euler's identity  $\langle \nabla f(g), g \rangle = -n f(g)/d$  for homogeneous functions yields:

$$-\frac{n}{d} \int_{\mathbf{G}} d\mathbf{x} = \langle g, \nabla f(g) \rangle = -\frac{n+d}{d} \int_{\mathbf{G}} g(\mathbf{x}) d\mathbf{x}.$$

□

### 3. THE $\ell_1$ -NORM FORMULATION

With  $d \in \mathbb{N}$  a fixed even integer and  $g \in \mathbf{Hom}_d$  written as

$$\mathbf{x} \mapsto g(\mathbf{x}) = \sum_{\alpha \in \mathbb{N}_d^n} g_\alpha \mathbf{x}^\alpha, \quad \mathbf{x} \in \mathbb{R}^m,$$

let  $\mathbf{B}_d \subset \mathbb{R}^n$  be the  $L_d$ -unit ball and  $\rho_d$  its Lebesgue volume, i.e.,

$$(3.1) \quad \mathbf{B}_d = \left\{ \mathbf{x} : \sum_{i=1}^n x_i^d \leq 1 \right\} \quad \text{and} \quad \rho_d := \text{vol}(\mathbf{B}_d) = \int_{\mathbf{B}_d} d\mathbf{x}.$$

Let  $\|g\|_1 := \sum_{\alpha \in \mathbb{N}_d^n} |g_\alpha|$  and consider the optimization problem  $\mathbf{P}_1$ :

$$(3.2) \quad \mathbf{P}_1 : \quad \inf_g \{ \|g\|_1 : f(g) = \rho_d; \quad g \in \mathbf{Hom}_d \}.$$

That is, among all degree- $d$  homogeneous polynomials  $g$  whose level set  $\mathbf{G}$  has same Lebesgue volume as the  $L_d$ -unit ball  $\mathbf{B}_d$ , one seeks the one which minimizes the  $\ell_1$ -norm of its coefficients.

In fact, since  $f(\lambda g) = \lambda^{-n/d} f(g)$  one may replace the constraint  $f(g) = \rho_d$  with the inequality constraint  $f(g) \leq \rho_d$  and (3.2) reads:

$$(3.3) \quad \mathbf{P}_1 : \quad \ell_1^* = \inf_g \{ \|g\|_1 : f(g) \leq \rho_d; \quad g \in \mathbf{Hom}_d \}.$$

**Theorem 3.1.** *Let  $d \geq 2$  be an even integer. The  $L_d$ -norm polynomial*

$$\mathbf{x} \mapsto g^*(\mathbf{x}) = \|\mathbf{x}\|_d^d \left( = \sum_{i=1}^n x_i^d \right),$$

*is the unique optimal solution of Problem  $\mathbf{P}_1$  in (3.3) and moreover,*

$$(3.4) \quad \text{vol}(\mathbf{B}_d) = \int_{\mathbf{G}^*} d\mathbf{x} = \frac{2^n \Gamma(1/d)^n}{n d^{n-1} \Gamma(n/d)}; \quad \int_{\mathbf{G}^*} x_i^d d\mathbf{x} = \frac{\text{vol}(\mathbf{B}_d)}{n+d} \quad i = 1, \dots, n.$$

*Proof.* Problem  $\mathbf{P}_1$  has an optimal solution  $g^* \in \mathbf{Hom}_d$ . Indeed let  $(g_k)$ ,  $k \in \mathbb{N}$ , be a minimizing sequence with  $\|g_k\|_1 \rightarrow \ell_1^* \geq 0$  as  $k \rightarrow \infty$ . Hence the sequence  $(g_n)$  is  $\ell_1$ -norm bounded and therefore there is a subsequence  $(k_t)$  and a polynomial  $g^* \in \mathbf{Hom}_d$  such that for every  $\alpha \in \mathbb{N}_d^n$ ,  $(g_{k_t})_\alpha \rightarrow g_\alpha^*$  as  $t \rightarrow \infty$ . Then of course one also obtains the pointwise  $g_{k_t}(\mathbf{x}) \rightarrow g^*(\mathbf{x})$ , as  $t \rightarrow \infty$ . Next, as  $f$  is nonnegative, by Fatou's Lemma

$$\begin{aligned} \rho_d \geq \liminf_{t \rightarrow \infty} f(g_{k_t}) &= \liminf_{t \rightarrow \infty} \frac{1}{\Gamma(1+n/d)} \int_{\mathbb{R}^n} \exp(-g_{k_t}(\mathbf{x})) d\mathbf{x} \\ &\geq \frac{1}{\Gamma(1+n/d)} \int_{\mathbb{R}^n} \liminf_{t \rightarrow \infty} \exp(-g_{k_t}(\mathbf{x})) d\mathbf{x} \\ &= \frac{1}{\Gamma(1+n/d)} \int_{\mathbb{R}^n} \exp(-g^*(\mathbf{x})) d\mathbf{x} \end{aligned}$$

which proves that  $g^*$  is feasible for  $\mathbf{P}_1$  and so is an optimal solution of  $\mathbf{P}_1$ .

Problem  $\mathbf{P}_1$  has the equivalent formulation:

$$\begin{aligned} \inf_{\lambda_\alpha, g_\alpha} \quad & \sum_{\alpha \in \mathbb{N}_d^n} \lambda_\alpha \\ \text{s.t.} \quad & \lambda_\alpha - g_\alpha \geq 0, \quad \forall \alpha \in \mathbb{N}_d^n \\ & \lambda_\alpha + g_\alpha \geq 0, \quad \forall \alpha \in \mathbb{N}_d^n \\ & f(g) \leq \rho_d; \quad \lambda_\alpha \geq 0, \quad \forall \alpha \in \mathbb{N}_d^n, \end{aligned}$$

which is a convex optimization problem for which Slater's condition holds. Hence at an optimal solution  $(g^*, \lambda)$ , the Karush-Kuhn-Tucker (KKT) optimality conditions

conditions read:

$$\begin{aligned}
1 - u_\alpha - v_\alpha - \psi_\alpha &= 0, \quad \forall \alpha \in \mathbb{N}_d^n \\
u_\alpha - v_\alpha + \theta \frac{\partial f(g^*)}{\partial g_\alpha} &= 0, \quad \forall \alpha \in \mathbb{N}_d^n \\
\lambda_\alpha, u_\alpha, v_\alpha, \psi_\alpha, \theta &\geq 0, \quad \forall \alpha \in \mathbb{N}_d^n \\
f(g^*) &\leq \rho_d \\
\lambda_\alpha \psi_\alpha = 0; u_\alpha (\lambda_\alpha - g_\alpha^*) &= 0, \quad \forall \alpha \in \mathbb{N}_d^n \\
\theta (1 - f(g^*)) = 0; v_\alpha (\lambda_\alpha + g_\alpha^*) &= 0, \quad \forall \alpha \in \mathbb{N}_d^n
\end{aligned}$$

for some dual variables  $(\mathbf{u}, \mathbf{v}, \psi, \theta)$ .

The meaning of the above optimality conditions is clear. Indeed at an optimal solution  $(g^*, \lambda)$  we must have  $\lambda_\alpha = |g_\alpha^*|$  for all  $\alpha$ . Moreover, from the complementarity conditions one also has  $u_\alpha v_\alpha = 0$  whenever  $g_\alpha^* \neq 0$ . In addition, from the two first equations, and the fact that  $1 = u_\alpha + v_\alpha = |u_\alpha - v_\alpha|$ ,  $\lambda_\alpha = |g_\alpha^*|$ , all moments  $\int_{\mathbf{G}^*} \mathbf{x}^\alpha d\mathbf{x}$  must be equal whenever  $g_\alpha^* \neq 0$ .

We next show that  $\mathbf{x} \mapsto g^*(\mathbf{x}) = \sum_{i=1}^n x_i^d$  is an optimal solution. Recall that  $\partial f(g)/\partial g_\alpha = \frac{n+d}{d} \int_{\mathbf{G}^*} \mathbf{x}^\alpha d\mathbf{x}$ . Choose

$$\theta := \frac{d}{n+d} \left( \int_{\mathbf{G}^*} x_1^d d\mathbf{x} \right)^{-1}; \quad u_\alpha = \theta \frac{n+d}{d} \int_{\mathbf{G}^*} \mathbf{x}^\alpha d\mathbf{x},$$

and  $v_\alpha = 0$ ,  $\lambda_\alpha = g_\alpha^*$ , for all  $\alpha \in \mathbb{N}_d^n$ . (Notice that  $u_\alpha = 1$  whenever  $\mathbf{x}^\alpha = x_i^d$  for some  $i$ .) Hence  $u_\alpha \geq 0$  for all  $\alpha \in \mathbb{N}_d^n$  (because  $u_\alpha = 0$  whenever some  $\alpha_i$  is odd), and let  $\psi_\alpha := 1 - u_\alpha$  for all  $\alpha \in \mathbb{N}_d^n$ . Observe that  $\psi_\alpha \geq 0$  because

$$\left| \int_{\mathbf{G}^*} \mathbf{x}^\alpha d\mathbf{x} \right| \leq \int_{\mathbf{G}^*} x_1^d d\mathbf{x}, \quad \forall \alpha \in \mathbb{N}_d^n.$$

Therefore,  $(g^*, \mathbf{u}, \mathbf{v}, \psi, \lambda, \theta)$  satisfies the (necessary) KKT-optimality conditions and as Slater's conditions holds and  $\mathbf{P}_1$  is convex, the KKT-optimality conditions are also sufficient. Hence we may conclude that  $g^*$  is an optimal solution of  $\mathbf{P}_1$ . Finally, observe that

$$\begin{aligned}
\ell_1^* = \|g^*\|_1 = n &= \sum_{\alpha \in \mathbb{N}_d^n} \lambda_\alpha = \sum_{\alpha \in \mathbb{N}_d^n} (u_\alpha - v_\alpha) g_\alpha^* \\
&= -\theta \langle \nabla f(g^*), g^* \rangle = \theta \frac{n}{d} f(g^*),
\end{aligned}$$

from which we deduce

$$\int_{\mathbf{G}^*} x_i^d d\mathbf{x} = \frac{1}{n+d} \int_{\mathbf{G}^*} d\mathbf{x},$$

for  $i = 1, \dots, n$ , which is (3.4). Finally, for the numerical value of  $\text{vol}(\mathbf{B}_d)$  in (3.4) see Lemma 7.1.

It remains to prove that  $g^*$  above is the unique optimal solution of  $\mathbf{P}_1$ . So suppose that  $\mathbf{P}_1$  has another optimal solution  $h \in \mathbf{Hom}_d$  (hence such that  $h \neq g^*$  and  $\|h\|_1 = \|g^*\|_1 = n$ ). As we have seen, necessarily  $f(h) = f(g^*) = \rho_d = \text{vol}(\mathbf{B}_d)$ . But then as  $\mathbf{P}_1$  is a convex optimization problem, any convex combination  $h_\lambda := \lambda h + (1 - \lambda)g^* \in \mathbf{Hom}_d$ ,  $\lambda \in (0, 1)$ , is also an optimal solution of  $\mathbf{P}$ . By strict convexity of  $f$ ,

$$f(h_\lambda) < \lambda f(h) + (1 - \lambda)f(g^*) = \rho_d.$$



But again by taking  $k^{-n/d}f(h_\lambda) = \rho_d$  (so that  $k < 1$ ) we exhibit another feasible solution  $\tilde{g} := k h_\lambda \in \mathbf{Hom}_d$  with smaller  $\ell_1$ -norm  $\|\tilde{g}\|_1 = k\|g^*\|_1$ , in contradiction with the fact that  $g^*$  is an optimal solution of  $\mathbf{P}$ . Hence  $g^*$  is the unique optimal solution of  $\mathbf{P}_1$ .  $\square$

**An alternative formulation.** We may also consider the alternative but equivalent formulation

$$(3.5) \quad \mathbf{P}'_1 : \quad \rho' = \inf_g \{ f(g) : \|g\|_1 \leq n; \quad g \in \mathbf{Hom}_d \}.$$

**Proposition 3.2.** *Let  $\mathbf{P}_1$  and  $\mathbf{P}'_1$  be as in (3.3) and (3.5), respectively. Then  $\mathbf{P}'_1$  and  $\mathbf{P}_1$  have same optimal value  $\rho_d$  and moreover the homogeneous polynomial  $\mathbf{x} \mapsto g^*(\mathbf{x}) = \sum_{i=1}^n x_i^d$  is the unique optimal solution of  $\mathbf{P}'_1$ .*

*Proof.* Again as  $f(\lambda g) = \lambda^{-n/d}f(g)$  we may only consider those  $h \in \mathbf{Hom}_d$  with  $\|h\|_1 = n$ . Suppose that  $h$  is an optimal solution of  $\mathbf{P}'_1$  with  $f(h) = \rho' < \rho_d$  and  $\|h\|_1 = n$ . Then take  $\tilde{g} = kh$  with  $k^{-n/(d)}\rho' = \rho_d$  so that  $k < 1$ . Then  $\|\tilde{g}\|_1 = k\|h\|_1 = kn < n$ . But this implies that  $\tilde{g}$  would be a better solution for  $\mathbf{P}_1$  than  $g^*$ , a contradiction. Therefore  $\rho' \geq \rho_d$  and in fact  $\rho' = \rho_d$  as  $g^*$  is feasible for  $\mathbf{P}'_1$  with  $f(g^*) = \rho_d$ . Next, observe that  $\mathbf{P}'_1$  is a convex optimization problem with a strictly convex objective function; hence an optimal solution is unique.  $\square$

Therefore problem  $\mathbf{P}_1$  has the equivalent formulation: *Among all homogeneous polynomials  $g \in \mathbf{Hom}_d$  with  $\|g\|_1 = 1$  which is the one with minimum volume?* By Theorem 2.1 the  $L_d$ -unit ball has minimum volume.

#### 4. THE $\ell_2$ -NORM FORMULATION

Let denote by  $\mathbf{x} \cdot \mathbf{y}$  the usual scalar product in  $\mathbb{R}^n$ . For every  $\alpha \in \mathbb{N}^n$  let  $c_\alpha := \frac{(\sum_i \alpha_i)!}{\alpha_1! \dots \alpha_n!}$ . Recall that  $\mathbb{N}_d^n := \{\alpha \in \mathbb{N}^n : |\alpha| = d\}$  and  $s_d = \binom{n-1+d}{d}$ . We now write

$$\mathbf{x} \mapsto p(\mathbf{x}) := \sum_{\alpha \in \mathbb{N}_d^n} c_\alpha p_\alpha \mathbf{x}^\alpha, \quad p \in \mathbf{Hom}_d,$$

for some vector  $\mathbf{p} = (p_\alpha) \in \mathbb{R}^{s_d}$ , and equip  $\mathbf{Hom}_d$  with the scalar product

$$\langle p, q \rangle_d := \sum_{\alpha \in \mathbb{N}_d^n} c_\alpha p_\alpha q_\alpha, \quad p, q \in \mathbf{Hom}_d,$$

with associated norm  $\|p\|_{2,d}^2 = \langle p, p \rangle_d$ . Next, denote by  $\mathcal{P}_d \subset \mathbf{Hom}_d$  the convex cone of homogeneous polynomials of degree  $d$  which are nonnegative, and let  $\mathcal{C}_d \subset \mathbf{Hom}_d$  be the convex cone of sums of  $d$ -powers of linear forms. Then  $\mathcal{C}_d$  is the dual cone of  $\mathcal{P}_d$ , i.e.,  $\mathcal{P}_d^* = \mathcal{C}_d$ ; see e.g. Reznick [15].

As in §3, let  $\rho_d = \text{vol}(\mathbf{B}_d)$  and consider the following optimization problem:

$$(4.1) \quad \mathbf{P}_2 : \quad \ell_2^* = \inf_g \{ \|g\|_{2,d}^2 : f(g) \leq \rho_d; \quad g \in \mathbf{Hom}_d \},$$

a (weighted)  $\ell_2$ -norm analogue of  $\mathbf{P}_1$  in (3.2). In view of Theorem 2.1, problem  $\mathbf{P}_2$  is a convex optimization problem.

**Theorem 4.1.** *Problem  $\mathbf{P}_2$  in (4.1) has a unique optimal solution  $g^* \in \mathbf{Hom}_d$  whose vector of coefficients  $\mathbf{g}^* = (g_\alpha^*) \in \mathbb{R}^{s_d}$  satisfies:*

$$(4.2) \quad g_\alpha^* = \ell_2^* \frac{n+d}{n} \cdot \frac{\int_{\mathbf{G}^*} \mathbf{x}^\alpha d\mathbf{x}}{\int_{\mathbf{G}^*} d\mathbf{x}}, \quad \forall \alpha \in \mathbb{N}_d^n,$$

where  $\mathbf{G}^* = \{\mathbf{x} : g^*(\mathbf{x}) \leq 1\}$  and  $\text{vol}(\mathbf{G}^*) = \rho_d$ .

Therefore,  $\mathbf{g}^* = (g_\alpha^*)$ ,  $\alpha \in \mathbb{N}_d^n$ , is an element of  $\mathcal{C}_d = \mathcal{P}_d^*$ . More precisely:

$$(4.3) \quad \mathbf{x} \mapsto g^*(\mathbf{x}) = \ell_2^* \frac{n+d}{n} \cdot \frac{\int_{\mathbf{G}^*} (\mathbf{z} \cdot \mathbf{x})^d d\mathbf{z}}{\int_{\mathbf{G}^*} d\mathbf{z}}$$

and in fact there exist  $\mathbf{z}_i \in \mathbb{R}^n$ ,  $\theta_i > 0$ ,  $i = 1, \dots, s$  with  $s \leq \binom{n-1+d}{d} + 1$ , such that

$$(4.4) \quad \mathbf{x} \mapsto g^*(\mathbf{x}) = \ell_2^* \frac{n+d}{n} \sum_{i=1}^s \theta_i (\mathbf{z}^i \cdot \mathbf{x})^d.$$

*Proof.* That  $\mathbf{P}_2$  has an optimal solution follows exactly with same arguments as for  $\mathbf{P}_1$ . Moreover Slater's condition also holds for  $\mathbf{P}_2$ . Hence by the KKT-optimality conditions, there exists  $\lambda^* \geq 0$  such that

$$2g_\alpha^* c_\alpha = -\lambda^* \frac{\partial f(g^*)}{\partial g_\alpha} = \lambda^* \frac{n+d}{d} \int_{\mathbf{G}^*} c_\alpha \mathbf{x}^\alpha d\mathbf{x}, \quad \forall \alpha \in \mathbb{N}_d^n.$$

Therefore, multiplying each side with  $g_\alpha^*$  and summing up yields

$$2\|g^*\|_{2,d}^2 = 2\ell_2^* = -\lambda^* \langle \nabla f(g^*), g^* \rangle = \lambda^* \frac{n}{d} f(g^*) = \lambda^* \frac{n}{d} \rho_d.$$

Hence  $\lambda^* = 4\ell_2^* d / (n\rho_d)$  and  $g_\alpha^* = \ell_2^* \frac{n+d}{n\rho_d} \int_{\mathbf{G}^*} \mathbf{x}^\alpha d\mathbf{x}$ , from which (4.2) follows.

But then (4.2) means that  $(g_\alpha^*) \in \mathcal{P}_d^* = \mathcal{C}_d$ , which yields (4.3). To get (4.4) we use a generalization of Tchakaloff's theorem described in [10], Anastassiou [1] and Kemperman [9].  $\square$

So both optimal solutions  $g_1^*$  of  $\mathbf{P}_1$  and  $g_2^*$  of  $\mathbf{P}_2$  are sums of  $d$ -powers of linear forms. But  $g_2^*$  does not have the parsimony property as shown below.

**Corollary 4.2.** *Let  $g^* \in \mathbf{Hom}_d$  be the optimal solution of  $\mathbf{P}_2$ . Then all its coefficients  $g_\alpha^*$  with  $\alpha = 2\beta \in \mathbb{N}_d^n$  for some  $\beta \in \mathbb{N}_{d/2}^n$ , are strictly positive. Hence if  $d \geq 4$  the optimal solution  $\mathbf{x} \mapsto g_1^*(\mathbf{x}) = \|\mathbf{x}\|_d^d$  of  $\mathbf{P}_1$  cannot be an optimal solution.*

*Proof.* From the characterization (4.2), every coefficient  $g_{2\beta}^*$  with  $2|\beta| = d$  must be positive. Hence the optimal solution  $g_1^*$  of problem  $\mathbf{P}_1$  cannot be an optimal solution of  $\mathbf{P}_2$ . Moreover there are  $\binom{n+d/2-1}{d/2}$  such coefficients.  $\square$

Corollary 4.2 states that the optimal solution of  $\mathbf{P}_2$  does not have a parsimony property as it has at least  $\binom{n+d/2-1}{d/2}$  non zero coefficients!

The only case where the optimal solution of  $\mathbf{P}_1$  also solves  $\mathbf{P}_2$  is the quadratic case  $d = 2$ . Indeed straightforward computation shows that (4.2) is satisfied by the polynomials  $g_1^*$  of Theorem 2.1.

**Example 1.** Let  $n = 2$  and  $d = 4$ . By symmetry we may guess that the optimal solution  $g^* \in \mathbf{Hom}_d$  is of the form:

$$\mathbf{x} \mapsto g^*(\mathbf{x}) = g_{40}(x_1^4 + x_2^4) + 6g_{22}x_1^2x_2^2,$$

with  $\ell_2^* = 2g_{40}^* + g_{22}^*$  and

$$g_{40} = 3\ell_2^* \frac{\int_{\mathbf{G}^*} x_1^4 d\mathbf{x}}{\int_{\mathbf{G}^*} d\mathbf{x}}; \quad g_{22} = 3\ell_2^* \frac{\int_{\mathbf{G}^*} x_1^2 x_2^2 d\mathbf{x}}{\int_{\mathbf{G}^*} d\mathbf{x}}.$$

Observe that by homogeneity, an optimal solution  $g^*$  of  $\mathbf{P}_2$  is also optimal when we replace  $\rho_d$  with any constant  $a$ ; only the optimal value changes and the characterization (4.2) remains the same with the new optimal value  $\ell_2^*$ . After several numerical trials we conjecture that

$$\mathbf{x} \mapsto g^*(\mathbf{x}) \approx x_1^4 + x_2^4 + 2x_1^2x_2^2 = (x_1^2 + x_2^2)^2,$$

i.e.  $\mathbf{g}^* = (1, 0, 1/3, 0, 1)$ , is an optimal solution. But then observe that

$$\mathbf{G}^* = \{\mathbf{x} : g_2^*(\mathbf{x}) \leq 1\} = \{\mathbf{x} : (x_1^2 + x_2^2)^2 \leq 1\} = \mathbf{B}_2!$$

That is,  $\mathbf{G}^*$  is another representation of the unit sphere  $\mathbf{B}_2$  by homogeneous polynomials of degree 4 instead of quadratics!

$$\int_{\mathbf{G}^*} d\mathbf{x} \approx 3.1415926; \quad \int_{\mathbf{G}^*} x_1^4 d\mathbf{x} \approx 0.392699; \quad \int_{\mathbf{G}^*} x_1^2 x_2^2 d\mathbf{x} \approx 0.130899.$$

With  $a := \int_{\mathbf{G}^*} d\mathbf{x}$ , (4.2) yields (up to  $10^{-8}$ )

$$3\ell_2^* \cdot \frac{\int_{\mathbf{G}^*} x_1^4 d\mathbf{x}}{\int_{\mathbf{G}^*} d\mathbf{x}} = 1 = g_{40}^*; \quad 3\ell_2^* \cdot \frac{\int_{\mathbf{G}^*} x_1^2 x_2^2 d\mathbf{x}}{\int_{\mathbf{G}^*} d\mathbf{x}} = \frac{1}{3} = g_{22}^*.$$

Observe also that

$$(x_1^2 + x_2^2)^2 = \frac{1}{6}(x_1 + x_2)^4 + \frac{1}{6}(x_1 - x_2)^4 + \frac{2}{3}(x_1^4 + x_2^4),$$

i.e., a sum of 4-powers of linear forms as predicted by Theorem 4.1.

In fact we have:

**Theorem 4.3.** *If  $d = 4$  then for every  $n$  the optimal solution of  $\mathbf{P}_2$  is  $g^*(\mathbf{x}) = (\sum_{i=1}^n x_i^2)^2$  whose level set  $\mathbf{G}$  is the unit ball  $\mathbf{B}_2$ .*

*Proof.* Let  $\mathbf{x} \mapsto g^*(\mathbf{x}) := (\sum_{i=1}^n x_i^2)^2$ . It is enough to prove that (4.2) holds (as by homogeneity (4.2) still holds when one replaces  $g$  with  $\lambda g$  for any  $\lambda > 0$ .) Since  $g^*(\mathbf{x}) = \sum_i x_i^4 + 2\sum_{i < j} x_i^2 x_j^2$ , we have  $\ell_2^* = (n + \frac{n(n-1)}{2} \cdot 6(\frac{2}{6})^2) = n(n+2)/3$ . Moreover

$$A := \int_{\mathbf{B}_2} x_1^4 d\mathbf{x} \left( \int_{\mathbf{B}_2} d\mathbf{x} \right)^{-1} = \frac{3}{(n+4)(n+2)}$$

Therefore with  $d = 4$ ,

$$\ell_2^* \frac{n+d}{n} A = \frac{n(n+2)}{3} \frac{n+4}{n} \frac{3}{(n+2)(n+4)} = 1 = g_{40}^*.$$

Similarly one has

$$B := \int_{\mathbf{B}_2} x_1^2 x_2^2 d\mathbf{x} \left( \int_{\mathbf{B}_2} d\mathbf{x} \right)^{-1} = \frac{1}{(n+4)(n+2)}$$

so that

$$\ell_2^* \frac{n+4}{n} B = \frac{n(n+2)}{3} \frac{n+4}{n} \frac{1}{(n+2)(n+4)} = \frac{1}{3} = g_{22}^*.$$

□

In other words, when  $d = 4$  the Euclidean unit ball  $\mathbf{B}_2 = \{\mathbf{x} : \sum_i x_i^2 \leq 1\}$  (which has the equivalent quartic representation  $\{\mathbf{x} : (\sum_{i=1}^n x_i^2)^2 \leq 1\}$ ) solves problem  $\mathbf{P}_2$ !

## 5. THE SOS FORMULATION

As we have seen that both optimal solutions of the  $\ell_1$ -norm and  $\ell_2$ -norm formulations are sums of  $d/2$  powers of linear forms, hence sums of squares (in short SOS). Therefore one may now restrict to homogeneous polynomials in  $\mathbf{Hom}_d$  that are SOS, i.e., polynomials of the form

$$\mathbf{x} \mapsto g_{\mathbf{Q}}(\mathbf{x}) = \mathbf{v}_{d/2}(\mathbf{x})^T \mathbf{Q} \mathbf{v}_{d/2}(\mathbf{x}),$$

where  $\mathbf{v}_{d/2}(\mathbf{x}) = (\mathbf{x}^\alpha)$ ,  $\alpha \in \mathbb{N}_{d/2}^n$ , and  $\mathbf{Q}$  is some real psd symmetric matrix ( $\mathbf{Q} \succeq 0$ ) of size  $s(d/2) = \binom{n-1+d/2}{d/2}$ . If we denote by  $\mathcal{S}_d$  the space of real symmetric matrices of size  $s(d)$ , there is not a one-to-one correspondence between  $g \in \mathbf{Hom}_d$  and  $\mathbf{Q} \in \mathcal{S}_d$  as several  $\mathbf{Q}$  may produce the same polynomial  $g_{\mathbf{Q}}$ .

Given  $0 \preceq \mathbf{Q} \in \mathcal{S}_d$ , denote by  $\mathbf{G}_{\mathbf{Q}}$  the sublevel set  $\{\mathbf{x} : g_{\mathbf{Q}}(\mathbf{x}) \leq 1\}$  associated with  $g_{\mathbf{Q}} \in \mathbf{Hom}_d$  and let  $f(\mathbf{Q}) := \text{vol}(\mathbf{G}_{\mathbf{Q}})$ . Observe that again  $f$  is positively homogeneous of degree  $-n/d$ .

So the natural analogue for  $\mathbf{Q}$  of the  $\ell_1$ -norm  $\|g\|_1$  for  $g \in \mathbf{Hom}_d$  is now the nuclear norm of  $\mathbf{Q}$  which as  $\mathbf{Q} \succeq 0$  reduces to  $\langle \mathbf{I}, \mathbf{Q} \rangle = \text{trace}(\mathbf{Q})$ . It is well-known that optimizing the nuclear norm on convex problems with matrices induce a parsimony effect, namely an optimal solution will generally have a small rank. In our context,  $\mathbf{Q}$  having a small rank means that  $g_{\mathbf{Q}}$  can be written a sum of a small number of squares. However, when expanded in the monomial basis,  $g_{\mathbf{Q}}$  may have many non-zero coefficients and so its  $\ell_1$ -norm  $\|g_{\mathbf{Q}}\|_1$  may not be small.

So in the same spirit as for the  $\ell_1$ - and  $\ell_2$ -norm, we now consider the optimization problem:

$$\mathbf{P}_3 : \quad \ell_3^* = \inf_{\mathbf{Q} \in \mathcal{S}_d} \{ \langle \mathbf{I}, \mathbf{Q} \rangle : f(\mathbf{Q}) \leq \rho_d; \quad \mathbf{Q} \succeq 0 \},$$

and characterize its unique optimal solution  $\mathbf{Q}^*$ .

**Theorem 5.1.** *Problem  $\mathbf{P}_3$  has a unique optimal solution. Moreover,  $\mathbf{Q}^* \in \mathcal{S}_d$  is an optimal solution of  $\mathbf{P}_3$  if and only if  $f(\mathbf{Q}^*) = \rho_d$  and:*

$$(5.1) \quad \mathbf{I} \succeq \frac{(n+d) \langle \mathbf{I}, \mathbf{Q}^* \rangle}{n \rho_d} \int_{\mathbf{G}_{\mathbf{Q}^*}} \mathbf{v}_{d/2}(\mathbf{x}) \mathbf{v}_{d/2}(\mathbf{x})^T d\mathbf{x}$$

where

$$\mathbf{x} \mapsto g_{\mathbf{Q}^*}(\mathbf{x}) := \mathbf{v}_{d/2}(\mathbf{x})^T \mathbf{Q}^* \mathbf{v}_{d/2}(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n.$$

Moreover, the polynomial  $\mathbf{x} \mapsto g(\mathbf{x}) = \sum_{i=1}^n x_i^d$  cannot be solution of  $\mathbf{P}_3$ .

*Proof.* Let  $(\mathbf{Q}_k)$ ,  $k \in \mathbb{N}$ , be a minimizing sequence. As  $\sup_k \langle \mathbf{I}, \mathbf{Q}_k \rangle \leq \langle \mathbf{I}, \mathbf{Q}_1 \rangle$ , the sequence  $(\mathbf{Q}_k)$  is norm-bounded in  $\mathcal{S}_d$ . Therefore it has a converging subsequence  $\mathbf{Q}_{k_j} \rightarrow \mathbf{Q}^* \in \mathcal{S}_d$  with  $\mathbf{Q}^* \succeq 0$  and  $f(\mathbf{Q}^*) \leq \rho_d$  (by a simple continuity argument). Therefore  $\mathbf{Q}^*$  is an optimal solution of  $\mathbf{P}_3$  and again, uniqueness follows from the strict convexity and homogeneity of the function  $f$ . Moreover, Slater's condition obviously holds for  $\mathbf{P}_3$  which is a convex optimization problem. Then the KKT-optimality conditions read:

$$(5.2) \quad \mathbf{I} + \lambda \nabla f(\mathbf{Q}^*) = \Psi \succeq 0$$

$$(5.3) \quad \langle \mathbf{Q}^*, \Psi \rangle = 0$$

$$(5.4) \quad f(\mathbf{Q}^*) \leq \rho_d; \quad \lambda(f(\mathbf{Q}^*) - \rho_d) = 0,$$

for some dual variables  $(\lambda, \Psi) \in \mathbb{R}_+ \times \mathcal{S}_d$ . By homogeneity one must have  $f(\mathbf{Q}^*) = \rho_d$ . Again Euler's identity for homogeneous functions yields  $\langle \nabla f(\mathbf{Q}^*), \mathbf{Q}^* \rangle = -nf(\mathbf{Q}^*)/d$ . Therefore using (5.2) and (5.3) one obtains  $\ell_3^* = \langle \mathbf{I}, \mathbf{Q}^* \rangle = \lambda n \rho_d / d$ , that is,  $\lambda = d \ell_3^* / (n \rho_d)$ . Next combining  $\Psi \succeq 0$  with Theorem 2.1, one also gets

$$\mathbf{I} \succeq \frac{n+d}{d} \lambda \int_{\mathbf{G}_{\mathbf{Q}^*}} \mathbf{v}_{d/2}(\mathbf{x}) \mathbf{v}_{d/2}(\mathbf{x})^T d\mathbf{x},$$

or, equivalently

$$\mathbf{I} \succeq \frac{(n+d) \langle \mathbf{I}, \mathbf{Q}^* \rangle}{n \rho_d} \int_{\mathbf{G}_{\mathbf{Q}^*}} \mathbf{v}_{d/2}(\mathbf{x}) \mathbf{v}_{d/2}(\mathbf{x})^T d\mathbf{x},$$

which is (5.1). This proves the *only if part* in Theorem 5.1.

Conversely, assume that  $0 \preceq \mathbf{Q}^* \in \mathcal{S}_d$  satisfies  $f(\mathbf{Q}^*) = \rho_d$  and (5.1). Let

$$\lambda := d \langle \mathbf{I}, \mathbf{Q}^* \rangle / (n \rho_d); \quad \Psi := \mathbf{I} - \frac{(n+d) \langle \mathbf{I}, \mathbf{Q}^* \rangle}{n \rho_d} \int_{\mathbf{G}_{\mathbf{Q}^*}} \mathbf{v}_{d/2}(\mathbf{x}) \mathbf{v}_{d/2}(\mathbf{x})^T d\mathbf{x}.$$

Obviously  $\lambda \geq 0$ ,  $\Psi \succeq 0$ , and:

$$\begin{aligned} \langle \mathbf{Q}^*, \Psi \rangle &= \langle \mathbf{I}, \mathbf{Q}^* \rangle \left[ 1 - \frac{n+d}{n \rho_d} \int_{\mathbf{G}_{\mathbf{Q}^*}} \langle \mathbf{Q}^*, \mathbf{v}_{d/2}(\mathbf{x}) \mathbf{v}_{d/2}(\mathbf{x})^T \rangle d\mathbf{x} \right] \\ &= \langle \mathbf{I}, \mathbf{Q}^* \rangle \left[ 1 - \frac{n+d}{n \rho_d} \int_{\mathbf{G}_{\mathbf{Q}^*}} g_{\mathbf{Q}^*}(\mathbf{x}) d\mathbf{x} \right] \\ &= \langle \mathbf{I}, \mathbf{Q}^* \rangle \left[ 1 - \frac{n+d}{n \rho_d} \frac{n}{n+d} \int_{\mathbf{G}_{\mathbf{Q}^*}} d\mathbf{x} \right] \quad [\text{by Theorem 2.1}] \\ &= 0, \end{aligned}$$

which shows that the triplet  $(\mathbf{Q}^*, \lambda, \Psi)$  satisfy the KKT-optimality conditions (5.2)-(5.4). As Slater's condition holds for  $\mathbf{P}_3$ , (5.2)-(5.4) are sufficient for optimality, which concludes the *if part* of the proof.

We next prove that  $\mathbf{x} \mapsto g(\mathbf{x}) := \sum_{i=1}^n x_i^d$  (so that  $\mathbf{G} = \mathbf{B}_d$ ) cannot be the optimal solution of  $\mathbf{P}_3$ . Among all  $\mathbf{Q} \succeq 0$  such that  $\sum_i x_i^d = \mathbf{v}_{d/2}(\mathbf{x})^T \mathbf{Q} \mathbf{v}_{d/2}(\mathbf{x})$ , the one that minimizes  $\text{trace}(\mathbf{Q})$  is  $\mathbf{Q} = \mathbf{I}$  with  $\text{trace}(\mathbf{Q}) = n$ . By Lemma 7.1, observe that

$$\frac{\int_{\mathbf{G}} x_i^d d\mathbf{x}}{\int_{\mathbf{G}} d\mathbf{x}} = \frac{\int_{\mathbf{B}_d} x_i^d d\mathbf{x}}{\rho_d} = \frac{1}{n+d},$$

and so (5.1) cannot hold because for instance the north-west and south-east corner elements of the matrix

$$\begin{aligned} \mathbf{A} &:= \mathbf{I} - \frac{(n+d) \langle \mathbf{I}, \mathbf{Q} \rangle}{n \rho_d} \int_{\mathbf{G}} \mathbf{v}_{d/2}(\mathbf{x}) \mathbf{v}_{d/2}(\mathbf{x})^T d\mathbf{x} \\ &= \mathbf{I} - (n+d) \frac{\int_{\mathbf{G}} \mathbf{v}_{d/2}(\mathbf{x}) \mathbf{v}_{d/2}(\mathbf{x})^T d\mathbf{x}}{\int_{\mathbf{G}} d\mathbf{x}} \end{aligned}$$

vanish whereas the north-east and south-west corner elements are non-zero, in contradiction with  $\mathbf{A} \succeq 0$ .  $\square$

The fact that the  $L_d$ -unit ball is not an optimal solution of  $\mathbf{P}_3$  is not a surprise as the sparsity-induced norm  $\text{trace}(\mathbf{Q})$  (when  $\mathbf{Q} \succeq 0$ ) aims at find a polynomial  $g_{\mathbf{Q}} \in \mathbf{Hom}_d$  which can be written as a sum of squares with as few terms as possible in the sum. On the other hand, the sparsity-induced norm  $\|g\|_1$  aims at finding a polynomials  $g \in \mathbf{Hom}_d$  with as few monomials as possible when  $g$  is expanded in the monomial basis. These can be two conflicting criteria!

## 6. EXTENSION TO GENERALIZED POLYNOMIALS

In this section  $d$  is now a (positive) rational with  $L_d$ -unit ball  $\{\mathbf{x} : \sum_{i=1}^n |x_i|^d \leq 1\}$ . Even though the function  $\mathbf{x} \mapsto \sum_{i=1}^n |x_i|^d$  is a “generalized polynomial” and not a polynomial any more, it is still a nonnegative positively homogeneous of degree  $d$  for which Theorem 2.1(a) applies. On the other hand, the vector space of positively homogeneous functions of degree  $d$  is not finite-dimensional and so for optimization purposes we need define an appropriate finite-dimensional analogue of  $\mathbf{Hom}_d$ .

We will use the notation  $|\mathbf{x}| \in \mathbb{R}_+^n$  for the vector  $(|x_1|, \dots, |x_n|)$  and  $|\mathbf{x}|^\alpha$  for the generalized monomial  $|x_1|^{\alpha_1} \dots |x_n|^{\alpha_n}$ , whenever  $\alpha \in \mathbb{Q}_+^n$ .

**Definition 6.1.** Let  $0 < d \in \mathbb{Q}$ . Define the space  $\mathcal{C}_d$  as:

$$(6.1) \quad \mathcal{C}_d := \left\{ \sum_{\alpha \in \mathbb{Q}_+^n} g_\alpha |\mathbf{x}|^\alpha : g_\alpha \in \mathbb{R}; \quad |\alpha| := \sum_{i=1}^n \alpha_i = d \right\}$$

where only finitely many coefficients  $g_\alpha$  are non-zero. Then:

$$\|g\|_1 = \sum_{\alpha} |g_\alpha|; \quad \|g\|_2^2 = \sum_{\alpha} g_\alpha^2.$$

The space  $\mathcal{C}_d$  is a real infinite-dimensional vector space and each element of  $\mathcal{C}_d$  is a positively homogeneous functions of degree  $d$ .

**Definition 6.2.** With  $q \in \mathbb{N}$  let  $\mathbb{Z}_q^n \subset \mathbb{R}^n$  be the lattice  $\{\mathbf{z} \in \mathbb{R}^n : q\mathbf{z} \in \mathbb{Z}^n\}$ . With  $0 < d \in \mathbb{Z}_q$ , denote by  $\mathcal{N}_{dq}^n$  the finite set  $\{\alpha \in \mathbb{Z}_q^n : \alpha \geq 0; \sum_{i=1}^n \alpha_i = d\}$  of cardinality  $m(d, q)$  and by  $\mathbf{Hom}_d^q \subset \mathcal{C}_d$  the vector space of functions  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  defined by:

$$(6.2) \quad \mathbf{Hom}_d^q := \left\{ \sum_{\alpha \in \mathcal{N}_{dq}^n} g_\alpha |\mathbf{x}|^\alpha : (g_\alpha) \in \mathbb{R}^{m(d, q)} \right\},$$

which is a finite-dimensional vector space.

For instance with  $n = 2$  and  $d = 1$  one has  $\mathbf{B}_1 := \{ \mathbf{x} \in \mathbb{R}^n : \sum_{i=1}^n |x_i| \leq 1 \}$  and with  $0 < q \in \mathbb{N}$ ,

$$\mathcal{N}_{1q} = \left\{ \left( \frac{k}{q}, \frac{q-k}{q} \right) : k = 0, \dots, q \right\}; \quad m(d, q) = q + 1,$$

and  $g \in \mathbf{Hom}_d^q$  can be written as

$$\mathbf{x} \mapsto g(\mathbf{x}) = \sum_{k=0}^q g_k |x_1|^{\frac{k}{q}} |x_2|^{\frac{q-k}{q}}, \quad \mathbf{x} \in \mathbb{R}^2,$$

for some vector  $\mathbf{g} = (g_k) \in \mathbb{R}^{q+1}$ .

Obviously  $\mathbf{Hom}_d^q$  is a vector space of dimension  $m(d, q)$  and every function in  $\mathbf{Hom}_d^q$  is positively homogeneous of degree  $d$ . As we did in §3, with  $g \in \mathbf{Hom}_d^q$  is associated the level set  $\mathbf{G} := \{ \mathbf{x} : g(\mathbf{x}) \leq 1 \}$ , and whenever  $\mathbf{G}$  has finite volume let  $f(g) := \text{vol}(\mathbf{G})$ ,  $g \in \mathbf{Hom}_d^q$ . It follows that  $f$  is positively homogeneous of degree  $-n/d$ . Therefore (2.2)-(2.4) in Theorem 2.1 holds for  $f$ . Finally, let

$$\mathbf{x} \mapsto g^*(\mathbf{x}) := \sum_{i=1}^n |x_i|^d; \quad \mathbf{B}_d = \mathbf{G}^*; \quad \rho_d = \text{vol}(\mathbf{G}^*),$$

and consider the finite-dimensional optimization problem

$$(6.3) \quad \mathbf{P}_{1q} : \quad \ell_1^* = \inf_g \{ \|g\|_1 : f(g) \leq \rho_d; \quad g \in \mathbf{Hom}_d^q \}.$$

When  $d < 1$  the unit ball  $\mathbf{B}_d$  is not convex and is not associated with a norm as can be seen in Figure 2 where  $d = 1/2$ .

However, we have the following analogue of Theorem 3.1

**Theorem 6.3.** *Let  $0 < q \in \mathbb{N}$  and  $0 < d \in \mathbb{Z}_{q/2}$ . The polynomial*

$$(6.4) \quad \mathbf{x} \mapsto g^*(\mathbf{x}) = \sum_{i=1}^n |x_i|^d$$

*is the unique optimal solution of Problem  $\mathbf{P}_{1q}$  in (6.3) and moreover,*

$$(6.5) \quad \text{vol}(\mathbf{B}_d) = \frac{2^n \Gamma(1/d)^n}{n d^{n-1} \Gamma(n/d)}; \quad \int_{\mathbf{G}^*} |x_i|^d d\mathbf{x} = \frac{\text{vol}(\mathbf{B}_d)}{n+d} \quad i = 1, \dots, n.$$

*Proof.* The proof is almost a verbatim copy of that of Theorem 3.1 except that we now have to deal with generalized moments  $\int_{\mathbf{G}} |\mathbf{x}|^\alpha d\mathbf{x}$ ,  $\alpha \in \mathcal{N}_{dq}^n$  instead of standard monomial moments  $\int_{\mathbf{G}} \mathbf{x}^\alpha d\mathbf{x}$ ,  $\alpha \in \mathbb{N}^n$ . But the crucial fact that we exploit is that  $f$  is strictly convex and Theorem 2.1 holds for  $f$ . As in the proof of Theorem 3.1, to show that  $g^*$  in (6.4) satisfies the KKT-optimality conditions we only need prove that

$$\int_{\mathbf{G}^*} |\mathbf{x}|^\alpha d\mathbf{x} \leq \int_{\mathbf{G}^*} |x_1|^d d\mathbf{x}, \quad \forall \alpha \in \mathcal{N}_{dq}^n.$$

Define the Hankel-type moment matrix  $\mathbf{M}$  to be the real symmetric matrix with rows and columns indexed by  $\alpha \in \mathcal{N}_{d\frac{n}{2}}^n$  and with entries

$$\mathbf{M}(\alpha, \beta) := \int_{\mathbf{G}^*} |\mathbf{x}|^{(\alpha+\beta)/2} d\mathbf{x}, \quad \alpha, \beta \in \mathcal{N}_{d\frac{n}{2}}^n.$$

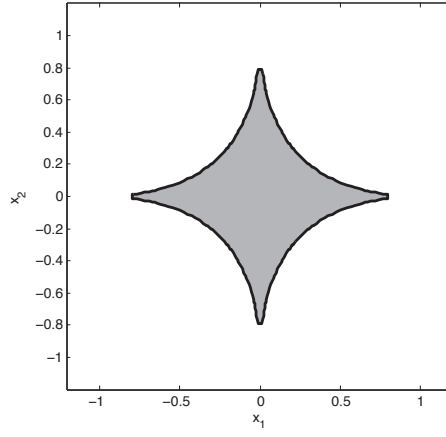


FIGURE 2. The  $L_{1/2}$ -unit ball  $\{ \mathbf{x} : \sqrt{|x_1|} + \sqrt{|x_2|} \leq 1 \}$

Equivalently, letting  $\mathbb{N}_{dq/2}^n = \{ \beta \in \mathbb{N}^n : \sum_i \beta_i = dq/2 \}$  and re-indexing rows and columns of  $\mathbf{M}$  with  $\tilde{\alpha} := q\alpha/2 \in \mathbb{N}_{dq/2}^n$ ,

$$\mathbf{M}(\tilde{\alpha}, \tilde{\beta}) := \int_{\mathbf{G}^*} (|\mathbf{x}|^{1/q})^{\tilde{\alpha} + \tilde{\beta}} d\mathbf{x} =: y_{\tilde{\alpha} + \tilde{\beta}}, \quad \tilde{\alpha}, \tilde{\beta} \in \mathbb{N}_{dq/2}^n.$$

Define  $\mathbf{y} = (y_{\tilde{\alpha}})$ ,  $\tilde{\alpha} \in \mathbb{N}_{dq}^n$ , and  $\tilde{X} = |\mathbf{x}|^{1/q}$ . Observe that from (6.2) one may write

$$\mathbf{Hom}_d^q := \left\{ \sum_{\tilde{\alpha} \in \mathbb{N}_{dq}^n} g_{\alpha} (|\mathbf{x}|^{\frac{1}{q}})^{\tilde{\alpha}} : (g_{\alpha}) \in \mathbb{R}^{m(d,q)} \right\}.$$



Let  $L_{\mathbf{y}} : \mathbf{Hom}_d^q \rightarrow \mathbb{R}_+$  be the linear mapping defined by

$$g \mapsto L_{\mathbf{y}}(g) := \sum_{\tilde{\alpha} \in \mathbb{N}_{dq}^n} g_{\tilde{\alpha}} y_{\tilde{\alpha}} = \sum_{\tilde{\alpha} \in \mathbb{N}_{dq}^n} g_{\tilde{\alpha}} \int_{\mathbf{G}^*} \tilde{X}^{\tilde{\alpha}} d\mathbf{x} = \int_{\mathbf{G}^*} g(\mathbf{x}) d\mathbf{x}.$$

By an adaptation of Lemma 4.3 in Lasserre and Netzer [12] to the present homogeneous context one has

$$|y_{\tilde{\alpha}}| \leq \sup_{i=1, \dots, n} L_{\mathbf{y}}(\tilde{X}_i^{dq}) = \sup_{i=1, \dots, n} \int_{\mathbf{G}^*} |x_i|^d d\mathbf{x} \left( = \int_{\mathbf{G}^*} |x_1|^d d\mathbf{x} \right), \quad \forall \tilde{\alpha} \in \mathbb{N}_{dq}^n.$$

Indeed in Lemma 4.3 of [12] one only uses the Hankel structure of the moment matrix  $\mathbf{M}$  and its positive definiteness. Therefore for every  $\alpha \in \mathcal{N}_{dq}^n$ ,

$$\int_{\mathbf{G}^*} |\mathbf{x}|^{\alpha} d\mathbf{x} = \int_{\mathbf{G}^*} \tilde{X}^{\tilde{\alpha}} d\mathbf{x} = y_{\tilde{\alpha}} \leq \int_{\mathbf{G}^*} |x_1|^d d\mathbf{x},$$

and so as in the proof of Theorem 3.1,  $g^*$  satisfies the KKT-optimality conditions.  $\square$

We obtain the following even more general extension of Theorem 3.1.

**Corollary 6.4.** *For every  $0 < d \in \mathbb{Q}$  the generalized polynomial*

$$\mathbf{x} \mapsto g^*(\mathbf{x}) = \sum_{i=1}^n |x_i|^d$$

*is the unique optimal solution of*

$$\mathbf{P}_1 : \quad \ell_1^* = \inf_g \{ \|g\|_1 : f(g) \leq \rho_d; \quad g \in \mathcal{C}_d \}.$$

*and (6.5) holds.*

*Proof.* Let  $0 < d \in \mathbb{Q}$  and suppose that there exists  $g \in \mathcal{C}_d$  such that  $\text{vol}(\mathbf{G}) = \rho_d$  and  $\|g\|_1 \leq n$ . Write  $d = p_0/q_0$  with  $0 < p_0, q_0 \in \mathbb{N}$ . For each non-zero coefficient  $g_{\alpha}$  one has  $\alpha_i = p_i(\alpha)/q_i(\alpha)$  for some integers  $0 < p_i(\alpha), q_i(\alpha)$ . Let  $q = 2q'$  with  $q' \in \mathbb{N}$  being the least common multiple (l.c.m.) of  $\{q_0, (q_i(\alpha)), i = 1, \dots, n, g_{\alpha} \neq 0\}$ . Then  $d \in \mathbb{Z}_{q/2}$  and  $g \in \mathbf{Hom}_d^q$ . Therefore by Theorem 6.3,  $\|g\|_1 > \|g^*\|_1 = n$  where  $g^*(\mathbf{x}) = \sum_{i=1}^n |x_i|^d$ , in contradiction with our assumption  $\|g\|_1 \leq n$ .  $\square$

Then again the parsimony property of the  $L_d$ -unit ball  $\mathbf{B}_d$  can be retrieved by minimizing the  $\ell_1$ -norm over all nonnegative generalized polynomials  $g \in \mathcal{C}_d$  whose associated ball unit ball  $\mathbf{G}$  has finite volume.

Next, concerning the  $\ell_2$ -norm, with  $0 < q \in \mathbb{N}$  an analogue of problem (4.1) now reads:

$$(6.6) \quad \mathbf{P}_{2q} : \quad \ell_2^* = \inf_g \{ \|g\|_2^2 : f(g) \leq \rho_d; \quad g \in \mathbf{Hom}_d^q \},$$

and we have the following analogue of Theorem 4.1:

**Theorem 6.5.** *With  $0 < q \in \mathbb{N}$  and  $0 < d \in \mathbb{Z}_q$ , Problem  $\mathbf{P}_{2q}$  in (6.6) has a unique optimal solution  $g^* \in \mathbf{Hom}_d^q$  whose vector of coefficients  $\mathbf{g}^* = (g_{\alpha}^*) \in \mathbb{R}^{m(d,q)}$*

satisfies:

$$(6.7) \quad g_\alpha^* = \ell_2^* \frac{n+d}{n} \cdot \frac{\int_{\mathbf{G}^*} |\mathbf{x}|^\alpha d\mathbf{x}}{\int_{\mathbf{G}^*} d\mathbf{x}}, \quad \forall \alpha \in \mathcal{N}_{dq},$$

where  $\mathbf{G}^* = \{\mathbf{x} : g^*(\mathbf{x}) \leq 1\}$  and  $\text{vol}(\mathbf{G}^*) = \rho_d$ .

We omit the proof as it is again a verbatim copy of that of Theorem 4.1. But in contrast to the case of polynomials in Theorem 4.1, in the optimal solution  $g^*$  of  $\mathbf{P}_{2q}$ , all coefficients  $(g_\alpha^*)$ ,  $\alpha \in \mathcal{N}_{dq}$ , are non-zero! This follows from (6.7) and the fact that all generalized moments  $\int_{\mathbf{G}^*} |\mathbf{x}|^\alpha d\mathbf{x}$ ,  $\alpha \in \mathcal{N}_{dq}$  are non-zero! For instance with  $d = 1/2$  and  $q = 8$ ,

$$0 < \int_{\mathbf{G}} |x_1|^{1/2} d\mathbf{x}; \quad 0 < \int_{\mathbf{G}} |x_1|^{1/8} |x_2|^{3/8} d\mathbf{x}; \quad 0 < \int_{\mathbf{G}} |x_1|^{1/4} |x_2|^{1/4} d\mathbf{x},$$

$$0 < \int_{\mathbf{G}} |x_1|^{3/8} |x_2|^{1/8} d\mathbf{x}; \quad 0 < \int_{\mathbf{G}} |x_2|^{1/2} d\mathbf{x}.$$

Hence the unique optimal solution  $g^*$  of  $\mathbf{P}_{2q}$  is not sparse at all. Even more, with fixed  $0 < d \in \mathbb{Q}$ , the larger is  $q$  the more complicated is  $g^*$ ! Therefore an analogue of Corollary 6.4 for the  $\ell_2$ -norm cannot exist.

## 7. APPENDIX

**Lemma 7.1.** *Let  $d$  be a positive real and let  $g : \mathbb{R}^n \rightarrow \mathbb{R}$  be the function:*

$$\mathbf{x} \mapsto g(\mathbf{x}) := \sum_{i=1}^n |x_i|^d; \quad \mathbf{G} := \{\mathbf{x} \in \mathbb{R}^n : g(\mathbf{x}) \leq 1\}.$$

*Then:*

$$(7.1) \quad \int_{\mathbf{G}} d\mathbf{x} = \frac{2^n}{n d^{n-1}} \frac{\Gamma(1/d)^n}{\Gamma(n/d)}; \quad \int_{\mathbf{G}} |x_i|^d d\mathbf{x} = \frac{2^n}{n(n+d) d^{n-1}} \frac{\Gamma(1/d)^n}{\Gamma(n/d)},$$

*for all  $i = 1, \dots, n$ .*

*Proof.* The function  $g$  is positively homogeneous of degree  $d$ . Observe that by Theorem 2.1,  $\int_{\mathbf{G}} d\mathbf{x} = \frac{1}{\Gamma(1 + n/d)} \int_{\mathbb{R}^n} \exp(-g(\mathbf{x})) d\mathbf{x}$ , and again by Theorem 2.1,

$$\int_{\mathbb{R}^n} \exp(-g(\mathbf{x})) d\mathbf{x} = \left( \int_{\mathbb{R}} \exp(-|t|^d) dt \right)^n = \left( \Gamma(1 + \frac{1}{d}) \int_{|t|^d \leq 1} dt \right)^n = \left( \frac{2\Gamma(1/d)}{d} \right)^n$$

where we have used the identity  $x\Gamma(x) = \Gamma(1+x)$ . This yields the result in the left of (7.1). Similarly, for every  $i = 1, \dots, n$ ,

$$\begin{aligned} \int_{\mathbb{R}^n} |x_i|^d \exp(-g(\mathbf{x})) d\mathbf{x} &= \left( \int_{\mathbb{R}} |t|^d \exp(-|t|^d) dt \right) \left( \int_{\mathbb{R}} \exp(-|t|^d) dt \right)^{n-1} \\ &= \left( \Gamma(1 + \frac{d+1}{d}) \int_{|t|^d \leq 1} |t|^d dt \right) \left( \Gamma(1 + \frac{1}{d}) \int_{|t|^d \leq 1} dt \right)^{n-1} \\ &= \left( \frac{2}{d} \Gamma((d+1)/d) \right) \left( \frac{2}{d} \Gamma(1/d) \right)^{n-1} \\ &= \frac{2^n}{d^{n+1}} \Gamma(1/d)^n. \end{aligned}$$

Therefore,

$$\int_{\mathbf{G}} |x_i|^d d\mathbf{x} = \frac{1}{\Gamma(1 + \frac{n+d}{d})} \int_{\mathbb{R}^n} |x_i|^d \exp(-g(\mathbf{x})) d\mathbf{x} = \frac{2^n}{n(n+d)d^{n-1}} \frac{\Gamma(1/d)^n}{\Gamma(n/d)}.$$

□

## REFERENCES

- [1] G.A. Anastassiou; *Moments in Probability and Approximation Theory*, Longman Scientific & Technical, UK, 1993.
- [2] F. Bach, R. Jenatton, J. Mairal, G. Obosinski. Optimization with sparsity-induced penalties, *Foundations and Trends in Machine Learning* **4**, issue **1**, 2013.
- [3] A. Beck, Y. Eldar. Sparsity constrained non linear optimization: Optimality conditions and algorithms, *SIAM J. Optim.* **23**, pp. 1480–1509, 2013.
- [4] E. Candès, J. Romberg, T. Tao. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information, *IEEE Trans. Inform. Theory* **52**, pp. 489–509, 2006.
- [5] D. L. Donoho. Compressed sensing, *IEEE Trans. Inform. Theory* **52**, pp. 1289–1306, 2006.
- [6] D. L. Donoho, M. Elad. Optimally sparse representation in general (non-orthogonal) dictionaries via  $\ell_1$  minimization, *Proc. Natl. Acad. Sci.* **100**, pp. 2197–2202, 2003.
- [7] E. Freitag and R. Busam. *Complex Analysis*, Second Edition, Springer-Verlag, Berlin, 2009.
- [8] D. Hilbert, S. Cohn-Vossen. *Geometry and the Imagination*, and Edition, Chelsea, 1952.
- [9] J.H.B. Kemperman. The general moment problem, a geometric approach, *Annals Math. Stat.* **39** (1968), pp. 93–122.
- [10] J.B. Lasserre. A generalization of Löwner-John’s ellipsoid theorem, *Math. Program.*, to appear. [arXiv:1302.1056](https://arxiv.org/abs/1302.1056)
- [11] J.B. Lasserre. A quick proof for the volume of  $n$ -balls. *Amer. Math. Monthly* **108**, pp. 768–769, 2001.
- [12] J.B. Lasserre and T. Netzer. SOS approximations of nonnegative polynomials via simple high degree perturbations *Math. Z.* **256**, pp. 99–112, 2007.
- [13] J.B. Lasserre and E.S. Zeron. Solving a class of multivariable integration problems via Laplace Lechniques, *Appl. Math. (Warsaw)* **28**, pp. 391–405, 2001.
- [14] B. Recht, M. Fazel, P. A. Parrilo. Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization, *SIAM Review* **52**, pp. 471–501, 2010.
- [15] B. Reznick. Sums of even powers of real linear forms, *Mem. Amer. Math. Soc.* **96**, No. 463, 1992.

LAAS-CNRS AND INSTITUTE OF MATHEMATICS, UNIVERSITY OF TOULOUSE, LAAS, 7 AVENUE DU COLONEL ROCHE, 31077 TOULOUSE CÉDEX 4, FRANCE, TEL: +33561336415  
*E-mail address:* [lasserre@laas.fr](mailto:lasserre@laas.fr)